

# bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted May 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 1

Alignment of Biotin Synthase Proteins from Barley (SID18), Corn (SID 20-24), Prickly Poppy (SID 28), Soybean (SID 28 and 30), and Wheat (SID 32)

SID18	TTTPS-----AVSP--SAAAPFRPALLAE--PAMLLLLARSLRSLRSPFASAVSAAPFSSVSAAAAEAAERAVRDGPRNDWTRP
SID20	MA-----LMLLARNLRLRPLPPLAA--AAAFSS--AAAEAAERAIIRDGPNDWSRP
SID22	MA-----LMLLARNLRLRPLPPLAA--AAAFSS--AAAEAAERAIIRDGPNDWSRP
SID24	MA-----LMLLARNLRLRPLPPLAA--AAAFSS--AAAEAAERAIIRDGPNDWSRP
SID26	-----MLKVQSL-RAR-----LRPLIFIS-----TFSSLSSSSSSAAAVQAERTIKEGPRNDWSRD
SID28	TKPNPKHYRCCLLSLCLYSQISHSFVSVLPNPFESKNMFLARPIFRA--PSLWALHSSYAYSSASAAAIQAERAIKEGPRNDWSRD
SID32	T-----MMLVRSVFRSQ-----LRPSV-SG-----GLQASCYSSLSAASAEARTIREGPRNDWSRD
1705463	MAT-----LRTSLRSLILLRSNTPKLAPIS-----SSVRLQVQKSRNYGTVSSVPPQATETSTSPSKDVYQEALNATE-PRSNWTRE
SID30	2995363M-----FT-----RTIRQQIRRSSALSLV-----RNNWTRE
6321725	M-----MSTIYRHLSTA--RPALTKYATNAA--VKSTTASSEASTLGALQYALS-----DEPSHSWTKS
SID18	EIQAIYDSPLLDLLFHGAQVHRNVHNFREVQQCTLLSIKTGGCSEDCSYCPQSSRYSTGLKAEKLMKDAVLEAAKKAKEAGSTRFCMGA
SID20	EIQAVYDSPLLDLLFHGAQ-----SSRYNTGLKAQKLMNKYAVLEAAKKAKEAGSTRFCMGA
SID22	EIQAVYDSPLLDLLFHGAQVHRNVHNFREVQQCTLLSIKTGGCSEDCSYCPQSSRYNTGLKAQKLMNKDAVLEAAKKAKEAGSTRFCMGA
SID24	EIQAVYDSPLLDLLFHGAQVHRNVHNFREVQQCTLLSIKTGGCSEDCSYCPQSSRYNTGLKAQKLMNKYAVLEAAKKAKEAGSTRFCMGA
SID26	EIKSVYDSPVLDLLFHAAQVHRHAHNFREVQQCTLLSVKTTGGCSEDCSYCPQSSRYDTGVKAQKLMNKDAVLEAAKKAKEAGSTRFCMGA
SID28	QVKSIVDSPILDLLFHGAQVHRHAHNFREVQQCTLLSIKTGGCSEDCSYCPQSSRYDTGVKGQRLMKNKEAVLQAAKKAKEAGSTRFCMGA
SID32	-----RDAVLEAAKKAKEAGSTRFCMGA
1705463	EIKSVYDSPLLDLLFHGAQVHRNVHNFREVQQCTLLSIKTGGCSEDCSYCPQSSRYSTGVKAQRLMSKDAVIDAAKKAKEAGSTRFCMGA
SID30	EIKAIYDKPLMELCWAGSLHRKFHPIPGA IQMCTLLNIKTGGCSEDCSYCAQSSRYQTGLKASKMWSVESVLAAR IAKONGSTRFCMGA
2995363	EIQKIYDTPLIDLIFRAASTHRKFPDPPKKVQVQCTLLSIKTGGCTEDCKYCAQSSRYNTGVKATKLMKIDEVLEKAKIAKAKGSTRFCMGS
6321725	QLKBEIYHTPLLELTHAAQLQHRKWHDP TKVQLCTLMNIKSGCSEDCCKYCAQSSRYNTGLKAEKMWKVDEVIKEAEAKRNGSTRFCMGA

[illegible]

FIGURE 1 (page 2 of 2)

SID18	AWRETI	GRKTNFNQ	ILEYVKD	IRGMGMEV	CC	TGLM	LEKQQA	BEELKK	AGLTAY	NHNLD	TSREYYP	NI	I	STRSYDD	R	Q	L	T	Q	H	V	R	E	A	G	I	S	V	C	
SID20	AWRETI	GRKSNFNQ	ILEYVKE	IRGMGMEV	CC	TGLM	IEKQQA	BEELKK	AGLTAY	NHNLD	TSREYYP	NI	I	STRSYDD	R	Q	L	T	Q	H	V	R	E	A	G	I	S	C		
SID22	AWRETI	GRKSNFNQ	ILEYVKE	IRGMGMEV	CC	TGLM	IEKQQA	BEELKK	AGLTAY	NHNLD	TSREYYP	NI	I	STRSYDD	R	Q	L	T	Q	H	V	R	E	A	G	I	S	C		
SID24	AWRETI	GRKSNFNQ	ILEYVKE	IRGMGMEV	CC	TGLM	IEKQQA	BEELKK	AGLTAY	NHNLD	TSREYYP	NI	I	STRSYDD	R	Q	L	T	Q	H	V	R	E	A	G	I	S	C		
SID26	AWRDT	VGRKTNFKQ	ILEYVKE	IRGMGMEV	CC	TGLM	IEKQQA	BEELKK	AGLTAY	NHNLD	TSREYYP	NI	I	STRSYD	R	E	L	E	T	Q	F	V	R	E	A	G	I	N	V	
SID28	AWRDT	LGRKTNFNQ	ILEYVKD	IRDMGMEV	CC	TGLM	LEKQQA	BEELKK	AGLTAY	NHNLD	TSREYYP	NI	I	TRTYD	R	E	L	Q	T	L	F	V	R	D	A	G	I	N	V	
SID32	AWRETI	GRKTNFNQ	ILEYVKD	IRGMGMEV	CC	TGLM	LEKQQA	BEELKK	AGLTAY	NHNLD	TSREYYP	NI	I	STRSYDD	R	Q	L	T	Q	H	V	R	E	A	G	I	S	V	C	
1705463	AWRDT	IGRKTNFSQ	ILEYI	KEIRGMGMEV	CC	TGLM	IEKQQA	BEELKK	AGLTAY	NHNLD	TSREYYP	PNV	I	STRSYD	R	E	L	T	L	S	H	V	R	D	A	G	I	N	V	
SID30	AWRDM	RGRKTNLKNV	KTMVSE	IRGMGMEV	CC	TGLM	IDAEQQA	BEELKK	AGLTAY	NHNVD	TSRDFY	PKVI	I	TRTYD	R	E	L	D	L	T	I	K	N	V	R	E	A	G	I	N
2995363	AWRDL	NGRNR	TFKN	ILEI	KEVRSMD	MEV	CC	TGLM	LENEQQA	BEELKK	DAGLTAY	NHNLD	TSREYYS	KI	I	STRTYD	R	E	L	N	T	I	D	N	L	R	K	A	G	L
6321725	AWRDM	KGRKSAMKR	IQEMV	TKVND	MGL	ET	CV	TGLM	VDQQA	KLK	DAGLTAY	NHN	ID	TSREHY	SKVI	I	TRTYD	R	Q	L	T	I	K	N	V	Q	E	S	G	I
SID18	SGGI	I	GLGEAE	EDRVGL	LLHT	LATL	PTH	ESVP	IN	AL	I	AVKGT	P	LQD	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
SID20	SGGI	I	GLGEAE	EDRVGL	LLHT	LATL	PTH	ESVP	IN	AL	I	AVKGT	P	LQD	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
SID22	SGGI	I	GLGEAE	EDRVGL	LLHT	LATL	PTH	ESVP	IN	AL	I	AVKGT	P	LQD	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
SID24	SGGI	I	GLGEAE	EDRVGL	LLHT	LATL	PTH	ESVP	IN	AL	I	AVKGT	P	LQD	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
SID26	SGGI	I	GLGEAE	EDRVGL	LLHT	LATL	PTH	ESVP	IN	AL	I	AVKGT	P	LQD	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
SID28	SGGI	I	GLGEAE	EDRVGL	LLHT	LATL	PTH	ESVP	IN	AL	I	AVKGT	P	LQD	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
SID32	SGGI	I	GLGEAE	EDRVGL	LLHT	LATL	PTH	ESVP	IN	AL	I	AVKGT	P	LQD	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
1705463	SGGI	I	GLGEAE	EDR	I	GL	LHT	LATL	PTH	ESVP	IN	AL	I	AVKGT	P	LQD	-	-	-	-	-	-	-	-	-	-	-	-	-	
SID30	TGGI	L	GLGENK	SDHI	GL	LE	T	VATL	PSH	ESFP	PN	ML	V	I	KGT	P	L	E	G	-	-	-	-	-	-	-	-	-	-	
2995363	SGGI	L	GLGEK	KH	D	R	V	GL	I	H	S	L	A	T	M	P	T	H	PS	V	P	F	N	L	L	V	I	P	G	T
6321725	TGGI	L	GLGE	SEDD	H	I	G	F	I	Y	T	L	S	N	M	S	P	H	ES	L										

FIGURE 2: Clone cdt2c.pk002.c17 Contains a 99 Nucleotide Deletion

SID19	CGACTGGAGCCGGCCCGAGATCCAGGCCGTCTACGACTCACCGCTCCTCGACCTCCTCTTT
SID21	CGACTGGAGCCGGCCCGAGATCCAGGCCGTCTACGACTCACCGCTCCTCGACCTCCTCTTT
SID19	TCACGGGGCTCAG-----
SID21	TCACGGGGCTCAGGTCACAGAAATGTCATAAATTCAGAGAAAGTGCAGCAATGCACACT
SID19	-----TCATCAAG
SID21	TCTTTCAATCAAGACTGGTGGATGCAGTGAAGATTGTTCTTACTGTCTCAGTCATCAAG
SID19	ATACAACACTGGATTGAAGGCCCAAAATTGATGAACAAAATATGCTGCTTTGGAAGCAGC
SID21	ATACAACACTGGATTGAAGGCCCAAAATTGATGAACAAAAGATGCTGCTTTGGAAGCAGC